PHYSICAL REVIEW A VOLUME 60, NUMBER 1 JULY 1999

Observing the spin of a free electron

B. M. Garraway[1] and S. Stenholm[2]
[1]*Sussex Centre for Optical and Atomic Physics, School of Chemistry, Physics, and Environmental Sciences, University of Sussex, Falmer, Brighton BN1 9QJ, England*
[2]*Physics Department, Royal Institute of Technology, Lindstedtsvägen 24, S-10044 Stockholm, Sweden*
(Received 8 October 1998)

Long ago, Bohr, Pauli, and Mott argued that it is not, in principle, possible to measure the spin components of a free electron. One can try to use a Stern-Gerlach type of device, but the finite size of the beam results in an uncertainty of the splitting force that is comparable with the gradient force. The result is that no definite spin measurement can be made. Recently there has been a revival of interest in this problem, and we will present our own analysis and quantum-mechanical wave-packet calculations which suggest that a spin measurement is possible for a careful choice of initial conditions. [S1050-2947(99)00207-3]

PACS number(s): 03.65.Bz, 03.75.−b, 14.60.Cd, 34.80.Nz

I. INTRODUCTION

The spin of the electron plays a central role in the explanation of atomic spectra. It turns out to endow the electron with a magnetic moment of magnitude $(e\hbar/2m)$, which carries the signature of its quantum origin in the factor $\hbar$. The magnetic moment of the orbital motion of the electron in an atom is of the same order, but it can be taken to a classical value by increasing the angular momentum quantum number. This has, indeed, been possible in Rydberg atoms, where nearly classical wave packets have been constructed and observed experimentally (see, for example, Ref. [1]). The spin of the electron, on the other hand, remains fixed; it cannot be increased, and thus the corresponding quantity has no classical limit. It is intrinsically a quantum property, which lead Bohr to the conclusion that it cannot be determined for a free electron. According to Bohr, all its observable effects are related to its role in the spectra of bound systems [2]. This point of view was put forward at the famous 1927 Como Conference by Bohr, but was not explicitly contained in the published version [3]. The reaction to the discussion is reviewed in the compilation of papers by Wheeler and Zurek [4].

Luckily the spirit of the Bohr argument has been preserved in the writings of two of his colleagues. Pauli reviewed the mathematics of the situation in [5], and Mott, who had contributed to the argument, published his version of the calculations both in Ref. [6] and in his text book [7]. In Ref. [8] Pauli explicitly stated that Bohr rejects the observability of the electron spin in situations where the concept of an electron trajectory is applicable. The conclusion is that the charge of the electron relates to its magnetic moment in such a manner that the separation of the spin components by the magnetic interaction is counteracted by the effect of the Lorentz force on the moving particle. The two effects are of the same order of magnitude, which can immediately be seen from the fact that the precession frequency (due to the magnetic moment) and the cyclotron frequency (bending the orbits) differ by only radiative corrections [9]. Ultimately this state of affairs derives from the origin of the electron spin in the Dirac theory, where only one magnetic coupling term occurs. Consequently, the quantum variable spin seems to be unmeasurable in a macroscopic apparatus of the type required by Bohr.

When the anomalous g-factor of the electron was measured by Dehmelt in a rf trap [10], it constituted a measurement of the electron spin. The issue created a controversy: was the spin trapped or free? According to Sir Rudolf Peierls, ''It is free in the sense intended by Bohr. This was one of the cases where Bohr was wrong'' [11]. Thus one may argue that the issue is settled; we can measure the spin of the ''free'' electron. However, the original argument of Bohr and Pauli still remains. It may well be worthwhile to investigate it in detail, to see how far it holds, and what possible ways there are to circumvent it. Such a discussion was recently initiated in Ref. [12], whose authors looked for experimental situations which allow a measurement of spin components. Their exploration was based on classical trajectories, and addressed Bohr's view as presented by Pauli. This is not to say that quantum mechanics does not play a role—in fact it does, as was made clear by Mott in his account of Bohr's argument [6] which used the uncertainty principle. Further, Bohr himself had second thoughts [13], and while he did not say that including quantum mechanics can effect a free-electron-spin splitting, he did say that he thinks that it is not possible to say it cannot. In the recent discussion of Ref. [12], similar arguments have been made [14] and further classical and quantum simulations of the Dirac equation have shown some incomplete spin separation [15].

In this paper we present a fully quantum-mechanical treatment as close to the Bohr-Pauli situation as is possible. We replace the classical trajectories by wave-packet motion, and implement the full quantum Hamiltonian for the two-level spin system. The formulation of the problem is identical to that chosen by Adler [16], who carried out a quantum-mechanical perturbation calculation. He agreed with our main conclusions that the experiment is, in principle, possible, and he essentially recommended the same parameter region that we utilize. However, due to the limitations of his perturbation approach, he could not quantify the exact degree of separation achievable. With modern computational techniques we are able to provide precise results of a nonperturbative version of the Bohr-Pauli experiment, and go beyond

it to suggest other possibilities.

One of us briefly discussed the argument by Bohr and Pauli [17], and it seems possible to manipulate the physical situation in such a way that the resolution of the spin measurement is improved. In addition, the original argument rests on a semiclassical interpretation of quantum mechanics, which omits essential features of the physics. Thus we consider it worthwhile to model the Stern-Gerlach experiment as a genuine particle experiment, where localized wave packets are launched into a magnetic field performing the roles assigned to it in the original argument. In this way, all quantum effects are taken into account, and, formulating the propagation in terms of two-component Pauli state vectors, we can also keep track of the proper vector character of the spin variable. All this can be treated as a time-dependent problem, and the possibility to separate the two spin components can be evaluated. We can decide the feasibility of the measurement, how far the Bohr-Pauli argument retains its validity, and which effects derive from the various quantum features of the problem.

For comparison, we set up the original argument in Sec. II. This is derived from the writings of Pauli and Mott; unfortunately, no details of the Bohr argument are available to us. In Sec. III we formulate the model, present its various features, and discuss its relation to the original argument. We find that the situation is far from trivial; many aspects enter into it, and the space of possible initial conditions is large and hard to survey.

In order to simplify the treatment, we throw away both the off-diagonality of the dipole interaction and the diamagnetic term, which eventually bends the orbit into cyclotron paths. This scalar model, presented in Sec. IV, can be solved analytically, and it contains all the essential features of the original argument. Thus we can utilize it to investigate the interplay between magnetic deflection and orbital bending, as long as these remain small enough. From these discussions, we determine how to look for parameter ranges where the spin measurement is most likely to succeed.

In Sec. V, we attack the full two-component quantum problem by solving the time-dependent Schrödinger equation numerically. The initial state is chosen according to the principles argued for within the scalar model. We present the results, and show to what extent the spins actually are split. As expected, the splitting is of the same order of magnitude as the widths of the individual components; to this extent the original argument is vindicated. However, there is a discernible split, and the resolution of the two components is a matter of numerical accuracy only. Spectral lines are resolved with less splitting in the laboratory. This, in our opinion, shows that the problem is of a practical character, not one of principle as claimed by Bohr.

The physical system we consider also shows some interesting features as a classical system; related problems have also been discussed in the chaos literature [18]. In order to check the consistency of our quantum calculations, we integrate the classical equations of motion for an ensemble of particles representing the position and momentum distribution of the initial quantum states. The results in Sec. VI reproduce those of the quantum calculations, but display a broader width. Thus the spin resolution is degraded, which shows the importance of performing a proper quantum analysis of the problem.

Then finally, there is the issue of principle: Is there a fundamental reason why magnetic fields cannot separate the spin components of a free electron? It appears that a negative answer has been found to this question already; there are the arguments we give above and there is the treatment in Ref. [12], but in Sec. VII we add to the discussion a configuration, which seems to allow an arbitrary spin separation, albeit in a situation differing from the beam experiment considered in the original argument. We utilize the magnetic field of the Stern-Gerlach experiment, and find an adiabatic method that separates the spin components in real space as far apart as we wish. There is, however, no beam, for the electron wave packet is allowed to glide in the magnetic field along a smooth trajectory. The feasibility of the experiment is not discussed, but as a demonstration of principle we regard the treatment as satisfactory. Finally, in Sec. VIII we summarize our discussion, and present a few conclusions and an outlook.

II. ORIGINAL ARGUMENT

We consider a beam of charged particles, with mass M and charge q. They propagate in the x direction with the linear momentum

$$p_x = \frac{2\pi\hbar}{\lambda_x}. \tag{1}$$

The beam is supposed to remain close to the x axis, with only small components in the y and z directions. If the particles have a magnetic moment μ and experience a magnetic field gradient in the z direction, the ensuing force is given by

$$F_z = \pm\mu\left(\frac{\partial B_z}{\partial z}\right), \tag{2}$$

where the sign depends on the direction of the magnetic moment, i.e., the spin. This force generates the splitting of the spin components, which is the basis for a Stern-Gerlach measurement.

However, the charge of the particle will also couple directly to the magnetic field. This gives a contribution to the z motion, from the Lorentz force, which competes with the measurement signal [Eq. (2)]. Because the magnetic field is sourceless, and here assumed to be independent of the x coordinate, we must have a component

$$\left(\frac{\partial B_y}{\partial y}\right) = -\left(\frac{\partial B_z}{\partial z}\right), \tag{3}$$

which causes a magnetic field proportional to the value of the location y. In quantum mechanics, an uncertainty ensues in the force, because the values of this location extend over some width Δy. Thus we obtain an uncertainty in the z force of the magnitude,

$$\Delta F_z = q\left(\frac{p_x}{M}\right)\left(\frac{\partial B_y}{\partial y}\right)\Delta y. \tag{4}$$

We demand that this is less than the spin-splitting due to the force [Eq. (2)] in order to resolve the components. But the magnetic moment is not an independent parameter. For a free electron it is of the form

$$\mu = \frac{e\hbar}{2m} \tag{5}$$

(within the accuracy we require here), where e and m are the charge and mass of an electron, respectively. For the magnetic moment of an atom, its order of magnitude is the same, only it is modified by the appropriate g factor.

Combining these results, we obtain

$$\frac{\Delta F_z}{F_z} = 4\pi\left(\frac{q}{e}\right)\left(\frac{m}{M}\right)\left(\frac{\Delta y}{\lambda_x}\right) \ll 1 \tag{6}$$

as a requirement for the observability of the spin. For an electron, all these ratios are unity or larger, and the conclusion is that electrons cannot be separated according to their spins by a Stern-Gerlach experiment. We proceed to make several comments on this calculation.

(1) For neutral atoms, $q=0$, and the experiment works; even for heavy ions, the factor (m/M) should make a separation possible [19].

(2) The whole argument is essentially classical; quantum mechanics enters only as an uncertainty. We will return to the discussion of this point below.

(3) The factor

$$\theta \sim \frac{\lambda_x}{\Delta y} \tag{7}$$

is a measure of the divergence angle of the beam; if we try to make $\Delta y \ll \lambda_x$, we no longer have a beam. However, the signal we want to resolve is in the z direction, and a large spreading in the y direction does not necessarily destroy the ability to distinguish between the two spin components if only condition (6) is satisfied.

(4) The Lorentz force is based on the longitudinal momentum p_x instead of the correct velocity; in the magnetic field, these are not equivalent entities. We will return to this question below.

The argument is based on the ratio between the forces, which we can easily verify to be correct. The momentum separation after the interaction time t_0 is

$$P_D = 2F_z t_0, \tag{8}$$

and the momentum spreading Δp_z is similarly obtained from Eq. (4). The spread in the z direction, Δz_0, is assumed not to change considerably from its initial value during the interaction. After the interaction has ceased, at times $t \gg t_0$, the spatial width over the separation becomes

$$\frac{\sqrt{\Delta z_0^2 + \left(\frac{\Delta p_z t}{m}\right)^2}}{P_D t/m} \Rightarrow \frac{\Delta F_z}{2F_z} \ll 1. \tag{9}$$

This shows that even for an arbitrarily large initial width Δz_0 the splitting eventually manifests itself in space, even though it may take a long time. Thus the necessary condition for the observability of the spin is that

$$\frac{\Delta p_z}{P_D} \ll 1 \tag{10}$$

after the end of the interaction time t_0. The experiment is thus expected to work best for the scenario depicted in Fig. 1. Here the spread in the z coordinate can be arbitrarily large, but according to Eq. (6) we want Δy to be as small as possible. In the following we try, utilizing a model of the physics outlined above, to explore to what extent the situation in Fig. 1 can be achieved within a quantum framework.

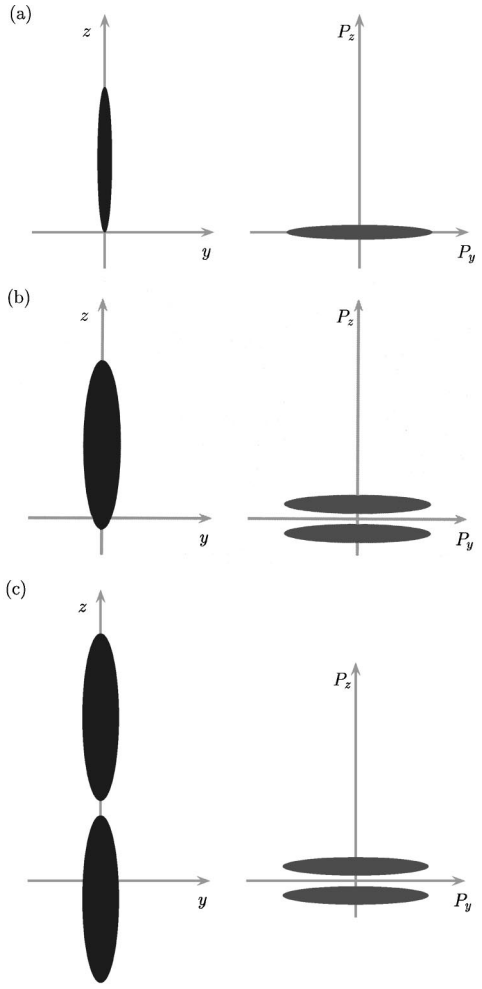

FIG. 1. This figure shows, in a schematic fashion, probability distributions in space (left) and momenta (right) for three different stages in the measurement process (a)–(c). In (a) we see the initial distributions for the electron. After the interaction with the magnetic field we move at time t_0 to (b), where a slight distortion of the spatial wave packet is seen, but without any significant spatial splitting. However, the splitting is seen in momentum space [(b), right]. The wave packets are then allowed to propagate in an interaction-free region. During this stage of the experiment, the splitting in momentum space will emerge spatially [(c), left].

III. MODEL

The beam of incident particles is taken to propagate in the x direction, and we choose the vector potential

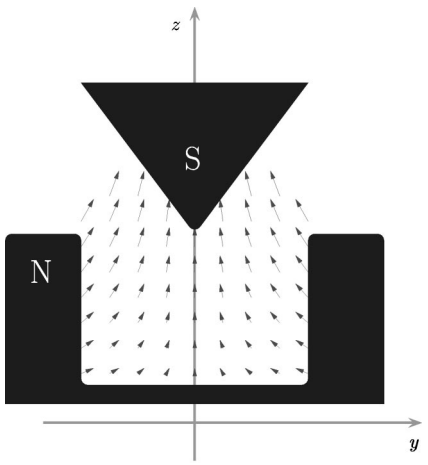

FIG. 2. Schematic illustration of magnets that will produce an approximation to the quadrupole magnetic field in Eq. (12). This resembles the conventional Stern-Gerlach arrangement (with the particle beam perpendicular to the section shown). An alternative approach is to generate the field with four long current-carrying wires (Fig. 4).

$$\mathbf{A}(\mathbf{r})=(-ayz,0,0), \tag{11}$$

where a is a parameter with dimensions of magnetic field over length. With this choice, $\hat{p}_x$ is a conserved quantity during the motion, and the operator can be replaced by a c number. The magnetic field becomes

$$\mathbf{B}(\mathbf{r})=\{0,-ay,az\}, \tag{12}$$

which obviously satisfies Eq. (3). The field is, in fact, a quadrupole field as indicated in Fig. 2; this corresponds to the upper part of the Stern-Gerlach apparatus as envisaged in the treatment above.

The magnetic coupling becomes

$$\mu\vec{\boldsymbol{\sigma}}\cdot\mathbf{B}=\mu a\begin{bmatrix} z & iy \\ -iy & -z \end{bmatrix}. \tag{13}$$

This suggests that in order to resolve the spin component splitting along the z axis, we need to insert the beam into the field at an offset z_0 in this direction; however to minimize the ill effects of the quantum off-diagonality, the initial wave packet should be as close to the z axis as is compatible with quantum theory. This conclusion agrees with the argument in Sec. II. We will quantify the effects of this off-diagonality in Sec. V C 1.

In the following, we introduce the mass of the electron m and its charge $q=-e$ into the equations. Then the full Schrödinger time evolution is given by

$$\left\{\frac{(p_x-eayz)^2}{2m}+\frac{p_y^2}{2m}+\frac{p_z^2}{2m}-\mu a\begin{bmatrix} z & iy \\ -iy & -z \end{bmatrix}\right\}\psi=i\hbar\frac{\partial\psi}{\partial t}, \tag{14}$$

where ψ is the two-component Pauli spinor.

We remove the constant term $p_x^2/2m$, but remember that the velocity in this direction must be obtained from the relation

$$v_x=\frac{p_x-eayz}{m} \tag{15}$$

whenever needed. In particular, this velocity carries the particles through the magnet, and it should not be allowed to change too much during the interaction. Again, we find that the experiment needs to be carried out near the origin in the y direction. The dynamics are then governed by the reduced Hamiltonian

$$H_R=\frac{p_y^2}{2m}+\frac{p_z^2}{2m}+V_1+V_2-\mu\vec{\sigma}\cdot\mathbf{B}, \tag{16}$$

where we have set

$$V_1=-\frac{p_x}{m}eayz \tag{17}$$

and

$$V_2=\frac{(eayz)^2}{2m}. \tag{18}$$

In order to estimate the contributions from the various terms in the Hamiltonian, we introduce a characteristic length l, such that the ratios (y/l) and (z/l) are of order unity. The dipole interaction term then becomes

$$\mu az\approx E_M\left(\frac{l}{\lambda_x}\right), \tag{19}$$

where we have introduced an energy characterizing the magnetic interaction

$$E_M\equiv\mu a\lambda_x=\frac{ea\hbar}{2m}\lambda_x. \tag{20}$$

The term linear in the vector potential is

$$|V_1|\sim 4\pi E_M\left(\frac{l}{\lambda_x}\right)^2, \tag{21}$$

and the second-order term becomes

$$V_2\sim E_M\left(\frac{l}{\lambda_x}\right)\left(\frac{eB_l l^2}{\hbar}\right), \tag{22}$$

where a characteristic size of the magnetic field is introduced:

$$B_l=al. \tag{23}$$

Equations (19), (21), and (22) offer a good starting point for a comparison of the main interactions: V_1 can be neglected with respect to the dipole force [Eq. (19)] when we have $l\ll\lambda_x$, as we already concluded in Sec. II for Δy. The orbit of the incoming beam particles is bending due to the cyclotron motion, and we may estimate the corresponding frequency to be

$$\omega_c=\frac{eB_l}{m}. \tag{24}$$

The time scale set by this motion should be compared with the spreading time T_s which doubles the width of a wave packet in free motion. If the size of the packet is originally of order l, the momentum uncertainty is estimated to be $\Delta p \sim \hbar/2l$, and the spreading time becomes [cf. Eq. (9) and Ref. [20]]

$$T_s = \frac{2ml^2}{\hbar}. \tag{25}$$

This bending of the beam trajectories plays no role in the original argument, and we will return to its influence below. Result (22) thus says that the interaction V_2 is negligible compared with the dipole interaction (19) when

$$\omega_c T_s \ll 1. \tag{26}$$

This is understandable: the cyclotron motion tries to bend the orbits of the particles; if this bending is small compared to the spreading of the wave packets, the sample can retain its character as a propagating beam. Finally, we note that in terms of ω_c and T_s the second order interaction V_2 can be written as

$$V_2 \sim \frac{1}{2} E_M \left(\frac{l}{\lambda_x} \right) \omega_c T_s . \tag{27}$$

IV. SCALAR MODEL

To solve the complete Schrödinger equation (14) is numerically possible, as we shall see in Sec. V, but a numerical approach offers no general indication about the parameter ranges favorable to the Stern-Gerlach experiment. In order to progress, we choose to treat a simplified model first, where we neglect (1) the term quadratic in the vector potential proportional to a^2, and (2) the off-diagonal terms in the dipole interaction.

We have discussed the conditions, under which we expect these to have negligible effects; their actual influence will be evaluated in the numerical work below. These two neglected terms play no role in the original argument because from Eq. (16) we obtain the equation of motion

$$\dot{p}_z = -\frac{\partial H_R}{\partial z} = \frac{eay}{m} p_x \pm \mu a. \tag{28}$$

The two contributions to the force in the z-direction are just the ones utilized in the argument in Sec. II.

We also carry out the scaling indicated above by introducing a length scale l, a time scale τ, and a corresponding momentum scale

$$\overline{p} = \frac{\hbar}{l}. \tag{29}$$

In the new totally dimensionless variables we have the Hamiltonian[1]

$$H_{\pm} = \tfrac{1}{2}(p_y^2 + p_z^2) - \Omega^2 yz \pm z, \tag{30}$$

where we have introduced a scaled wave number for the beam:

$$\Omega^2 = \frac{4\pi l}{\lambda_x}. \tag{31}$$

The scaling parameters achieving Eq. (30) obey the relations

$$\begin{aligned} l^3 &= \frac{2\hbar}{ea}, \quad l\tau = \frac{2m}{ea}, \\ \frac{l^2}{\tau} &= \frac{\hbar}{m}, \quad \frac{l}{\tau^2} = \frac{\mu a}{m}. \end{aligned} \tag{32}$$

In Eq. (30) all variables, including the scaled wave number Ω^2, are dimensionless quantities. We also see directly that the only way to decrease the ratio of the Lorentz force term to that of the dipole force is to make Ω small, i.e., make λ_x large. However, this will at some point destroy the beam quality, as discussed above.

In the following development, we wish to be able to keep track of the direction of the spin, and so we introduce a parameter $\kappa = \pm 1$, which gives the direction of the splitting force. We then carry out the coordinate transformation

$$\begin{aligned} \xi &= \frac{1}{\sqrt{2}}(y+z), \\ \eta &= \frac{1}{\sqrt{2}}(y-z). \end{aligned} \tag{33}$$

With these new coordinates, the Hamiltonian becomes

$$\begin{aligned} H = &\left[\frac{1}{2} p_\xi^2 - \frac{\Omega^2}{2} \left(\xi - \frac{\kappa}{\sqrt{2}\Omega^2} \right)^2 \right] \\ &+ \left[\frac{1}{2} p_\eta^2 + \frac{\Omega^2}{2} \left(\eta - \frac{\kappa}{\sqrt{2}\Omega^2} \right)^2 \right]. \end{aligned} \tag{34}$$

Thus we have reduced the solution to the case of two decoupled harmonic oscillators, albeit one is inverted.

In the Heisenberg representation, the operator solution to the dynamics under the Hamiltonian (34) is

$$\begin{aligned} \frac{\partial \hat{\xi}}{\partial t} &= \widehat{p_\xi} = \widehat{p_\xi^0} \cosh \Omega t + \Omega \left(\widehat{\xi^0} - \frac{\kappa}{\sqrt{2}\Omega^2} \right) \sinh \Omega t, \\ \frac{\partial \hat{\eta}}{\partial t} &= \widehat{p_\eta} = \widehat{p_\eta^0} \cos \Omega t - \Omega \left(\widehat{\eta^0} - \frac{\kappa}{\sqrt{2}\Omega^2} \right) \sin \Omega t. \end{aligned} \tag{35}$$

From these results we can solve for the momentum in the direction of interest

[1] We indicate the scaled dimensionless variables with the same symbols as their physical counterparts. We believe that, with proper care, we do not introduce any confusion by this.

$$\widehat{p_z}=\frac{1}{2}[\widehat{p_z^0}F_2(\Omega t)+\widehat{p_y^0}F_1(\Omega t)]+\frac{\Omega}{2}[\widehat{z^0}F_3(\Omega t)+\widehat{y^0}F_4(\Omega t)]$$

$$-\frac{\kappa}{2\Omega}F_4(\Omega t), \tag{36}$$

where we define

$$\begin{aligned} F_1(\Omega t)&=\cosh\Omega t-\cos\Omega t\rightarrow(\Omega t)^2,\\ F_2(\Omega t)&=\cosh\Omega t+\cos\Omega t\rightarrow 2+O((\Omega t)^4)\\ F_3(\Omega t)&=\sinh\Omega t-\sin\Omega t\rightarrow\frac{1}{3}(\Omega t)^3,\\ F_4(\Omega t)&=\sinh\Omega t+\sin\Omega t\rightarrow 2\Omega t. \end{aligned} \tag{37}$$

The limits are for $\Omega t\rightarrow 0$. Taking the expectation value of Eq. (36) over the initial state, we find the only operator to give a nonzero value to be the offset $\langle\widehat{z^0}\rangle=z_0$. Thus we find

$$\langle\widehat{p_z}(t)\rangle=\frac{\Omega z_0}{2}F_3(\Omega t)-\frac{\kappa}{2\Omega}F_4(\Omega t). \tag{38}$$

The terms proportional to $\kappa=\pm 1$ give the momentum splitting, which thus becomes

$$P_D=\frac{F_4(\Omega t)}{\Omega}. \tag{39}$$

We now define the dispersion in $\widehat{p_z}$ as

$$\sigma_z^2=\langle(\widehat{p_z}-\langle\widehat{p_z}\rangle)^2\rangle. \tag{40}$$

If we assume the initial dispersions to be uncorrelated, and define

$$\begin{aligned} \langle(\widehat{p_y^0})^2\rangle&=\Delta p_y^2,\quad \langle(\widehat{p_z^0})^2\rangle=\Delta p_z^2\\ \langle(\widehat{y^0})^2\rangle&=\Delta y^2,\quad \langle(\widehat{z^0}-z_0)^2\rangle=\Delta z^2, \end{aligned} \tag{41}$$

we obtain

$$\sigma_z^2=\tfrac{1}{4}[\Delta p_y^2F_1(\Omega t)^2+\Delta p_z^2F_2(\Omega t)^2+\Omega^2\Delta y^2F_4(\Omega t)^2+\Omega^2\Delta z^2F_3(\Omega t)^2]. \tag{42}$$

The resolution of the measurement is now, analogously with Eq. (9), given by

$$s^2(t)=\frac{\sigma_z^2}{P_D^2}=\frac{\Omega^2}{4}\left\{\Delta p_y^2\left(\frac{F_1(\Omega t)}{F_4(\Omega t)}\right)^2+\Delta p_z^2\left(\frac{F_2(\Omega t)}{F_4(\Omega t)}\right)^2+\Omega^2\Delta z^2\left(\frac{F_3(\Omega t)}{F_4(\Omega t)}\right)^2+\Omega^2\Delta y^2\right\}. \tag{43}$$

We cannot let $t\rightarrow 0$ here, because the term in Δp_z^2 must clearly diverge in that limit. For times such that $\Omega t\gtrsim 4$, all ratios $F_1(\Omega t)/F_4(\Omega t), F_2(\Omega t)/F_4(\Omega t), F_3(\Omega t)/F_4(\Omega t)$ are approximately unity, which means that we can omit these ratios and minimize the remaining terms separately.

The position uncertainty terms give

$$\frac{\Omega^4}{4}(\Delta y^2+\Delta z^2)\ll 1, \tag{44}$$

which we may rescale by using Eqs. (32) to obtain

$$4\pi^2\left(\frac{\Delta y^2+\Delta z^2}{\lambda_x^2}\right)\ll 1. \tag{45}$$

The left-hand side is found to be small exactly when the original argument gives a small ratio in Eq. (6), but Eq. (45) requires this condition to be valid isotropically in the y,z plane.

When the momentum terms are unscaled to physical units, using definitions (32), we find

$$\frac{\Omega^2}{4}(\Delta p_y^2+\Delta p_z^2)\Rightarrow 2\pi\left(\frac{\Delta p_y^2+\Delta p_z^2}{2m}\right)\left(\frac{1}{\mu a\lambda_x}\right)=2\pi\frac{\Delta E_{yz}}{E_M}\ll 1. \tag{46}$$

This shows that the energy associated with the momentum fluctuations in the y,z plane, ΔE_{yz}, must be less than the characteristic magnetic energy (20).

From Eqs. (44) and (46) we select the y components of position and momentum (or the corresponding z components) and multiply them together to obtain the combined condition

$$\left(\frac{\Delta y^2}{\lambda_x^2}\right)\left(\frac{\Delta p_y^2}{2m\mu a\lambda_x}\right)=\frac{(\Delta y\Delta p_y)^2}{2m\mu a\lambda_x^3}\ll\frac{1}{(2\pi)^3}. \tag{47}$$

However, the uncertainty is minimized if the product $\Delta y\Delta p_y=\hbar/2$, and consequently we have the condition

$$2\pi^3\frac{(\hbar^2/2m\lambda_x^2)}{\mu a\lambda_x}=\frac{\pi}{2}\left(\frac{E_x}{E_M}\right)\ll 1, \tag{48}$$

where $E_x=(2\pi\hbar)^2/2m\lambda_x^2$. Thus we require the kinetic energy along the beam, i.e., E_x, to be less than the magnetic interaction. This is the kinetic energy expressed in terms of the momentum p_x and not the kinematic energy in terms of the velocity. However, for small values of y, the deviations are small; cf. Eq. (15).

From Eq. (45) we also have

$$E_x=4\pi^2\left(\frac{\hbar^2}{2m\lambda_x^2}\right)\ll\frac{\hbar^2}{2m\Delta y^2}\sim\Delta E_{yz}. \tag{49}$$

With such large transverse fluctuations it seems difficult to claim that this describes a beam any more, and to that extent the original argument seems to be validated. We cannot separate the two spin directions into different well-defined

propagating beams. In the long-time limit, the parameters necessary to obtain a small value of s^2 destroy the beam quality because of the quantum nature of the motion.

However, quantum mechanics allows us one more attempt. We found that we could not use the long-time limit, because it is not compatible with the assumption of a propagating beam. In the short-time limit, s^2 diverges, but we can try to circumvent that.

We choose minimum uncertainty wave packets with $\Delta p_y \Delta y = \frac{1}{2}$ etc. (in scaled units). Then we can try to optimize the combined uncertainties in the y and z directions independently. We first use

$$\frac{\partial}{\partial \Delta y^2}\left[\frac{1}{4\Delta y^2}\left(\frac{F_1(\Omega t)}{F_4(\Omega t)}\right)^2+\Omega^2\Delta y^2\right]=0. \tag{50}$$

The result is

$$\frac{1}{4\Delta y^2}\left(\frac{F_1(\Omega t)}{F_4(\Omega t)}\right)^2+\Omega^2\Delta y^2 \Rightarrow \Omega\left(\frac{F_1(\Omega t)}{F_4(\Omega t)}\right). \tag{51}$$

For the z component we find

$$\frac{\partial}{\partial \Delta z^2}\left[\frac{1}{4\Delta z^2}\left(\frac{F_2(\Omega t)}{F_4(\Omega t)}\right)^2+\Omega^2\Delta z^2\left(\frac{F_3(\Omega t)}{F_4(\Omega t)}\right)^2\right]=0. \tag{52}$$

The result is

$$\frac{1}{4\Delta z^2}\left(\frac{F_2(\Omega t)}{F_4(\Omega t)}\right)^2+\Omega^2\Delta z^2\left(\frac{F_3(\Omega t)}{F_4(\Omega t)}\right)^2 \Rightarrow \Omega\left(\frac{F_2(\Omega t)F_3(\Omega t)}{F_4(\Omega t)^2}\right). \tag{53}$$

From Eq. (43) we now find

$$s^2=\frac{\Omega^3}{4}\left[\left(\frac{F_1(\Omega t)}{F_4(\Omega t)}\right)+\left(\frac{F_2(\Omega t)F_3(\Omega t)}{F_4(\Omega t)^2}\right)\right] =\frac{\Omega^3}{4}\left[\frac{\sinh 2\Omega t-\sin 2\Omega t}{(\sinh \Omega t+\sin \Omega t)^2}\right]. \tag{54}$$

In this form, however, a fortuitous cancellation appears. Using limits (37), we obtain

$$s^2 \Rightarrow \frac{\Omega^4 t}{6}. \tag{55}$$

We thus find that, for short enough durations $t=t_0$ of the interaction, we can make the resolution parameter s^2 as small as we like. This, however, implies that we choose an initial wave packet which is very unisotropic. The uncertainties should satisfy the relations

$$\Delta y^2=\frac{t_0}{4}, \quad \Delta p_y^2=\frac{1}{t_0},$$

$$\Delta z^2=\frac{3}{\Omega^4 t_0^3}, \quad \Delta p_z^2=\frac{\Omega^4 t_0^3}{12}. \tag{56}$$

For small enough t_0, these relations suggest the situation shown in Fig. 1.

We still need to rescale the relations to obtain physical quantities. We find

$$s^2 \Rightarrow \frac{(4\pi)^2 t_0}{6\lambda_x^2}\left(\frac{l^2}{\tau}\right)=\frac{4}{3}\left(\frac{(2\pi)^2\hbar^2}{2m\lambda_x^2}\right)\frac{t_0}{\hbar} \ll 1. \tag{57}$$

This says that the duration of the interaction has to satisfy an energy relation of the type

$$E_x t_0 \ll \tfrac{3}{4}\hbar. \tag{58}$$

This is not an uncertainty relation, but it indicates that the interaction time has to be short compared with the time characteristic of the free evolution of the wave packet, which takes place with the energy E_x.

Rescaling the uncertainties, we find the results

$$\Delta y^2=\frac{\hbar t_0}{4m} \ll \frac{3\hbar^2}{16mE_x}. \tag{59}$$

This again implies

$$\Delta E_y=\left(\frac{\hbar^2}{8m\Delta y^2}\right) \gg \frac{2}{3}E_x, \tag{60}$$

which seems to destroy the quality of the beam. However, the actual beam propagates with the velocity v_x, and this aspect will be discussed separately. As a consequence of Eq. (60), the requirement $\Delta y \ll \lambda_x$ from Eq. (6) re-emerges. However, in contrast to requirement (45), no similar restriction on Δz is found. This is supposed to be large, in order to allow a good resolution in p_z. From Eqs. (56) and (57) we find

$$\Delta z^2=\left(\frac{3}{4\pi^2}\right)\left(\frac{\lambda_x^2}{t_0^3}\right)\left(\frac{m^3}{e^2a^2\hbar}\right) \gg \frac{8}{9}(2\pi)^4\left(\frac{\hbar}{ea\lambda_x^2}\right)^2 =\frac{8}{9}\lambda_x^2\left(\frac{E_x}{E_M}\right)^2. \tag{61}$$

It does not seem necessary to impose any relationship of type (48) on the ratio (E_x/E_M) in the present, optimized scheme.

We have, however, one more check to make on the consistency of the scheme. As we saw in Sec. III, the interaction V_2, which has been neglected here, concerns the cyclotron motion in the magnetic field. This tries to bend the orbit back onto itself, which will destroy the beam character. We have to ask how much bending would occur in the time t_0, used to separate the linear momenta of the spin directions. This discussion has to be added because we have not included that term in the present considerations.

The main magnetic force in the y,z plane is due to the beam offset in the z direction; near the origin the magnetic field is zero. We calculate the corresponding cyclotron frequency to be

$$\omega_c = \frac{2\pi}{T_c} = \frac{eaz_o}{m}. \tag{62}$$

The magnetic field is in the z direction, and thus causes a bending primarily in the y direction. Because the kinetic energy is conserved, magnetic forces do no work, and the motion in the x direction will be affected as well. This will eventually stop the beam from propagating through the experimental setup. This will happen in spite of the fact that p_x is conserved; the motion forward is to be described by the velocity v_x; cf. Eq. (15). Some reduction in v_x may not be a bad thing. We can understand this because, in the original Bohr-Pauli argument, the velocity p_x/m should be replaced by v_x. If v_x is smaller than p_x/m, because of the magnetic field, we may find that the fluctuations ΔF_z are smaller than expected, thus allowing a resolution of spin.

In order to minimize the bending due to the cyclotron motion we have to require

$$t_0\omega_c \ll 1. \tag{63}$$

This gives the relation

$$t_0\left(\frac{eaz_0}{m}\right) = \frac{2t_0E_M}{\hbar}\left(\frac{z_0}{\lambda_x}\right) \ll 1. \tag{64}$$

In analogy with Eq. (58), we thus obtain the condition

$$E_M t_0 \ll \hbar\left(\frac{\lambda_x}{2z_0}\right). \tag{65}$$

If we use the offset $z_0 \gtrsim \Delta z$, condition (65) might suggest using $\Delta z \ll \lambda_x$, which, however, from Eq. (61) implies $E_x \ll E_M$, in which case Eq. (65) may be difficult to satisfy. On the other hand, if it is possible to use $z_0 \ll \lambda_x \ll \Delta z$, then we can take $E_M \approx E_x$, and retain the validity of Eqs. (58) and (65) simultaneously. These considerations suggest that it is possible to achieve a small resolution parameter s^2 in an interaction time which does not cause a considerable bending of the beam.

V. WAVE-PACKET DYNAMICS

A. Formulation of wave-packet problem

In order to test our ideas concerning the visibility of the spin splitting, we have performed a numerical integration of the full, time-dependent Schrödinger problem as given in Eq. (14). It is necessary to scale the problem, as in Sec. IV [Eq. (32)], so that we numerically integrate the Schrödinger equation

$$\left\{-\frac{1}{2}\left(\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+\frac{1}{2}(\Omega^2/2-2yz)^2-\begin{bmatrix} z & iy \\ -iy & -z\end{bmatrix}\right\}\psi = i\frac{\partial\psi}{\partial t}, \tag{66}$$

using, for example, the split-operator fast Fourier transform method. (For a summary of this and other integration methods, see Ref. [20].) Typically, the numerical integration is performed by discretizing the wave function on a two-dimensional (2D) grid of points (e.g., 256×256).

For the numerical calculations we retain the terms that were dropped in Sec. IV, i.e., we retain the $(yz)^2$ term arising from $\mathbf{A}^2$, and the iy term from the spin interaction. As before, we include the term proportional to yz, i.e., the cross term in the expansion of $(\Omega^2/2-2yz)^2$, which corresponds to the Ω^2 term in Eq. (30). After the scaling of the problem, we are left with a single free parameter, Ω, which determines the scaled momentum in the beam (x) direction. However, the parameters of the initial state, and especially its spatial distribution are still free to be chosen.

As our initial state we choose a Gaussian wave function of the form

$$\psi(y,z,t=0) = \frac{1}{\sqrt{2\pi\Delta y\Delta z}}\exp\left[-\frac{(y-y_0)^2}{4\Delta y^2}-\frac{(z-z_0)^2}{4\Delta z^2}\right]\times\begin{bmatrix}a_1\\ a_2\end{bmatrix}, \tag{67}$$

where Δy and Δz are the uncertainties in position as described in Sec. II and Fig. 1. The state is centered on (y_0,z_0), and has no net momentum, although, since it is a minimum uncertainty state, there are fluctuations in momentum given by $\Delta p_y \Delta y = \frac{1}{2}$, etc. The amplitudes a_1 and a_2 determine the probabilities in the σ_z basis; these are the probabilities we wish to measure by means of a clear splitting of the wave packets. For the purposes of demonstration we wish to discern two different trajectories of wave packets, and so we set both amplitudes equal to $1/\sqrt{2}$. There is, of course, no initial spatial difference between the two states in any basis.

B. Uncharged particle

In the absence of charge, the scaled Hamiltonian of Eq. (66) reduces to

$$H_{q=0} = -\frac{1}{2}\left(\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)-\begin{bmatrix} z & iy \\ -iy & -z\end{bmatrix}. \tag{68}$$

For this case Fig. 3 shows an example of an ordinary Stern-Gerlach-type splitting of spin-wave packets. The initial state is shown (for one of the levels) in Fig. 3(a). The wave packet is placed on the z axis of the coordinate system at a distance from the origin. This ensures that the spin splitting takes place in the σ_z basis; if the packet were placed on the y axis, at a distance from the origin, the splitting would be manifest in the σ_y basis.

The width of the initial wave packet has been chosen so that, in the time it takes for the two components to display a clear spatial separation, there is little spreading of the wave packets. In Figs. 3(b) and 3(c) we see the two wave-packet components at a later time. They are clearly at different physical locations, and thus a simple spatial separation of the wave packets will correspond to a spin separation of the components. The splitting is seen here as something that develops as a function of time because we do not view the translational motion in the longitudinal (x) direction. In prac-

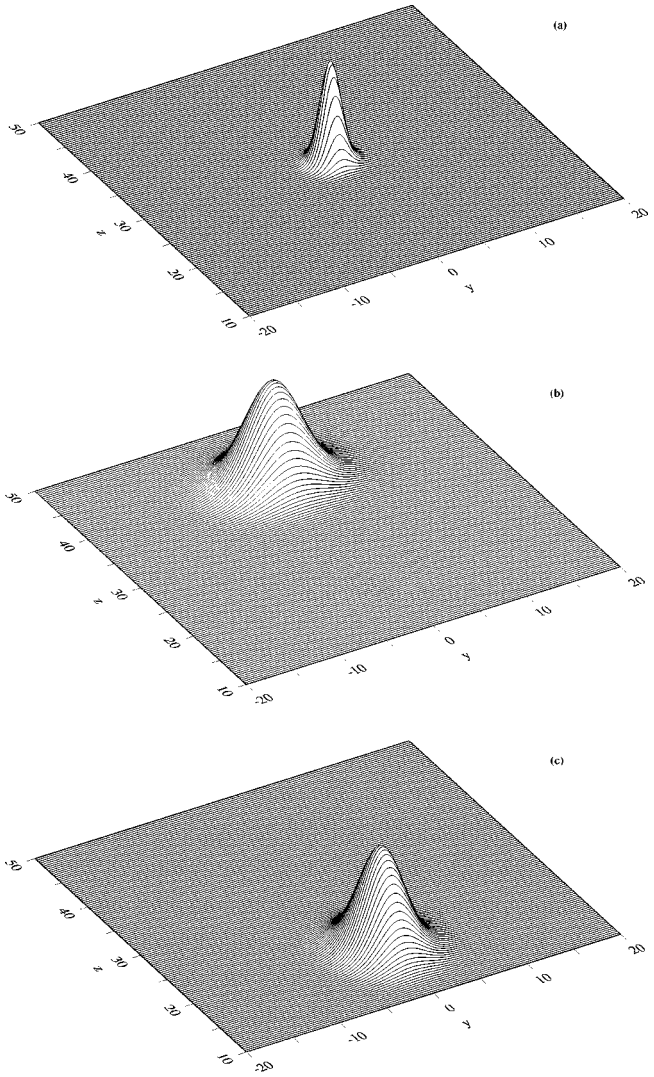

FIG. 3. Splitting of wave-packet spin components for a particle with no charge. In (a) we see the initial wave function as a function of y and z, which is located at $y_0=0$, $5z_0=30$, with a width of unity. In (b) and (c) we see the two wave-packet spin components after a time $t=4.5$ (in scaled units). The two components are clearly separated. They are now much broader, because of wave-packet spreading. The two components in (b) and (c) have slightly different shapes because the wave packet approaching the origin is affected by the increasing curvature of the adiabatic surfaces.

tice the beam splits so that our wave-packet pictures correspond to sections, in the y-z plane, perpendicular to the direction of the beam.

C. Considerations for a charged particle

When we include the charge of the particle, we must return to the integration of the wave equation (66) which contains the additional potential-energy term

$$V_L=\tfrac{1}{2}(\Omega^2/2-2yz)^2. \tag{69}$$

If we now place a wave packet on the z axis, this potential will cause a sideways deflection of the two components. The effect of the term linear in yz was fully taken into account in Sec. IV. For very large amplitudes, the motion is complicated, but in part consists of oscillations in the potential well (69). These oscillations change their frequency in time because the cyclotron frequency depends on field strength, which in turn depends on the location of the electron.

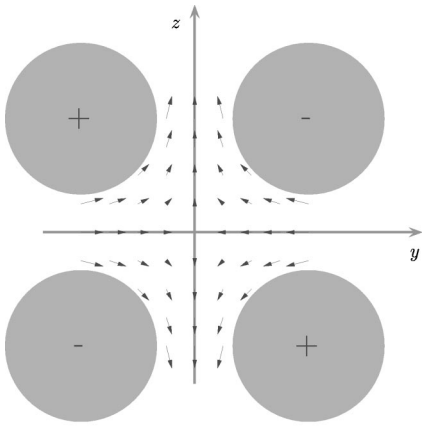

FIG. 4. Magnetic quadrupole field produced by four wires (seen in cross section). This allows the initial wave packet to be placed in the center of the field pattern.

We will next find constraints on the initial wave packet by considering some simple, desirable, qualities of the measurement process. In the following we will assume that our initial state is a minimum uncertainty packet which is unisotropic such that $\Delta z \gg \Delta y$. The initial Gaussian packet is chosen to be centered on the origin of the quadrupole field (i.e., $z_0=0$); this is justified in Sec. V C 2. At this point we cannot have a magnetic field produced by the traditional arrangement of Fig. 2 because the material of the magnet would get in the way. However, we can easily make a suitable arrangement of magnetic poles, or four current carrying wires, as shown in Fig. 4. Then the wave packet is straightforwardly centered on the origin.

1. Angular width of the wave packet

For a point particle placed on the z axis, we can see from Eq. (66) that the Stern-Gerlach coupling is initially of the form $-\sigma_z z$, which provides the force trying to split the spin components. However, if a point particle is placed on the y axis instead, the effective Stern-Gerlach coupling is initially of the form $-\sigma_y y$, which would effect a different kind of wave-packet splitting. Thus for a wave packet which is spread out over the y,z space, it is desirable that it should be tightly localized on the z axis of the y-z plane for the type of measurement that we want to make; if it is too spread out, we are unable to resolve the measurement of the z components of the spin. This requirement is the same as trying to reduce the off-diagonality, as we mentioned earlier in connection with Eq. (13). In order to characterize the spreading of the initial wave packet, we will consider the angular width of the peak in the initial angular probability distribution; this width will be denoted as $\Delta\theta$. Again, we consider a centralized wave packet for reasons we consider below. With an unisotropic wave packet, such as the one illustrated in Fig. 1(a), we can keep the angular spread low by reducing the width in the y direction. This remains true even if the wave packet is centered on the origin.

The angular distribution of the Gaussian wave packet (67) is easily obtained by radial integration of the probability distribution. That is, with $y_0=z_0=0$,

$$P(\theta)=\int_0^\infty dr\, r|\psi(r\cos\theta,r\sin\theta)|^2$$

$$=\frac{1}{2\pi}\frac{\Delta y\Delta z}{\Delta y^2+(\Delta z^2-\Delta y^2)\cos^2\theta}, \tag{70}$$

where we have $\tan\theta=z/y$. This is a two-peaked distribution, and we will use the width of one of the peaks to determine $\Delta\theta$. Because the wave packet has been chosen to be centered on the origin, distribution (70) is symmetric about $\theta=\pi$ and the two peaks have equal heights and equal widths. For $\Delta z>\Delta y$, one peak in this distribution is at $\theta=\pi/2$, with another peak at $\theta=-\pi/2$, and the full width at half maximum height gives a convenient measure of the width of either of the peaks of the distribution, i.e.,

$$\sin(\Delta\theta/2)=\frac{1}{\sqrt{\left(\frac{\Delta z}{\Delta y}\right)^2-1}}, \tag{71}$$

which for $\Delta z\gg\Delta y$ leads to the approximate expression

$$\Delta\theta\sim 2\frac{\Delta y}{\Delta z}. \tag{72}$$

The requirement that this angular width be small, $\Delta\theta\ll 1$, then simply leads to

$$\frac{\Delta y}{\Delta z}\ll\frac{1}{2}, \tag{73}$$

which clearly means that the wave packet must have a high degree of spatial anisotropy. This is completely consistent with the conclusion of Sec. IV.

2. Cyclotron period

In this section, in line with Eq. (63) of Sec. IV, we wish to choose sufficiently short interaction times t_0 for the cyclotron motion to be neglected. However, because the magnetic field varies, the cyclotron frequency ω_c depends on position in the magnet, and therefore it varies across the spatial distribution of the initial wave packet. At a distance r from the origin the cyclotron frequency is

$$\omega_c=\frac{ea}{m}r. \tag{74}$$

In a uniform magnetic field, the electron would orbit in circles, and after a time t it would have traversed an angle of $\omega_c t$. In our problem, the magnetic field is *not* uniform, but nevertheless we can introduce an angle $\phi=\omega_c t$ as an estimate of the bending that takes place for very short times (i.e., for the regime we aim for in Sec. IV). For the quadrupole field [Eq. (12)], and in the scaled units, Eqs. (32), the cyclotron frequency is simply $\omega_c=2r$, where r is the distance from the origin. Then at any point in space, the angle ϕ has the value $2rt_0$ at the end of the interaction. However, we have a wave packet that is distributed over space, with different magnetic-field strengths acting on its different parts. Because the unisotropic wave packet extends more in the z direction than the y direction, we take the phase angle to be $2zt_0$ at points along the length of the wave packet. The average value will be zero (because the bending is in different directions for positive and negative z), but the spread of the angle ϕ will be given approximately by

$$\Delta\phi=2\Delta z t_0. \tag{75}$$

However, according to Eqs. (56) the optimized time is

$$t_0=4\Delta y^2, \tag{76}$$

and hence the cyclotron phase becomes

$$\Delta\phi=8\Delta y^2\Delta z. \tag{77}$$

3. Initial cyclotron motion

The velocity in the x direction of any part of the wave packet is given by Eq. (15), which in scaled units is

$$v_x=\Omega^2/2-2yz, \tag{78}$$

as may be expected from Eq. (69). Because the generalized momentum in the x direction is a constant of the motion, Ω is fixed and v_x depends on y and z. Now we can see from Eq. (78) that if $yz<(\Omega/2)^2$ the velocity v_x is positive. However, if $yz>(\Omega/2)^2$ the velocity v_x is negative at that point, resulting in that part of the wave packet moving in a backward direction. The line defined by

$$yz=(\Omega/2)^2 \tag{79}$$

divides the forward- and backward-going regions, and we will call this the *cyclotron line*. This line is also the minimum of the Lorentz potential (69).

In pursuing the idea of a beamlike device, we will regard it as undesirable to have major parts of the wave packet moving backward, and to quantify this we will determine the amount, P_c, of the wave packet moving backward for the initial state. We use scaled polar coordinates $\{r,\theta\}$ defined by

$$y=r\Delta y\cos\theta, \tag{80}$$

$$z=r\Delta z\sin\theta. \tag{81}$$

Then the cyclotron line obeys the equation

$$r^2\sin 2\theta=\beta^2, \tag{82}$$

where

$$\beta^2=\frac{\Omega^2}{2\Delta y\Delta z}. \tag{83}$$

If we now integrate the wave packet over the two outer regions (i.e., the two large r regions in the first and third quadrants of the y-z plane), we obtain

$$P_c=\frac{1}{\pi}\int_0^{\pi/2}d\theta\int_{\beta/\sqrt{\sin 2\theta}}^{\infty}drre^{-r^2/2}=\frac{1}{2}\mathrm{erfc}\left(\frac{\Omega}{2\sqrt{\Delta y\Delta z}}\right). \tag{84}$$

This shows the expected limits that $P_c\to 0$ for small wave packets ($\Delta y\Delta z\to 0$), and $P_c\to\frac{1}{2}$ for very large initial wave packets ($\Delta y\Delta z\to\infty$) (when half the initial wave packet moves backward).

4. Summary of conditions

By combining Eq. (55) and Eqs. (56), we can express the parameter s, defined in Eq. (43), in terms of the widths of the wave packet:

$$s=\frac{1}{4\sqrt{2}\Delta y^2\Delta z}. \tag{85}$$

The parameter combination $\Delta y^2\Delta z$ can be regarded as not only determining s, but also some of our other essential parameters, which thus allows us to re-express them in terms of the parameter s. The cyclotron phase angle (77) can be written as

$$\Delta\phi=\frac{\sqrt{2}}{s}. \tag{86}$$

Furthermore, from Eqs. (56) we have

$$\Omega^2=\frac{\sqrt{3}}{t_0^{3/2}\Delta z}=\frac{\sqrt{3}}{2^3\Delta z\Delta y^3}. \tag{87}$$

Thus with Eq. (85) the square of the argument of the error function in Eq. (84) becomes

$$\frac{\Omega^2}{4\Delta y\Delta z}=\frac{\sqrt{3}}{2^5\Delta z^2\Delta y^4}=\sqrt{3}s^2, \tag{88}$$

and we can express it as

$$P_c=\tfrac{1}{2}\mathrm{erfc}(3^{1/4}s). \tag{89}$$

In order to have spin separation, we require that

$$s^2<1, \tag{90}$$

but we see in Eqs. (89) and (86) that, as s is reduced, P_c will approach one-half, i.e., the initial packet moves backward as much as forward, and, the phase angle $\Delta\phi$ increases rapidly (starting at 1.42 rad when $s=1$). It is necessary to make a compromise and thus, for example, we adopt a modest value of $s=0.8$; then $P_c\sim 0.07$ and $\Delta\phi\sim 102°$.

The phase angle $\Delta\phi$ is high because the extreme edge of the wave packet will move in a backward direction at the end of the interaction time. However, the bulk of the wave packet is in weaker fields and will not be affected so strongly. The main obstacle to reducing s in the optimized scheme appears to be the effects of the increasing phase angle $\Delta\phi$.

For the numerical computations we choose a narrow wave packet with an angular width of $\Delta\theta=0.01$, which determines the ratio of the uncertainties Δy and Δz, and together with $s=0.8$ and Eqs. (56) we obtain $\Delta y\sim 0.103$ and $\Delta z\sim 20.7$.

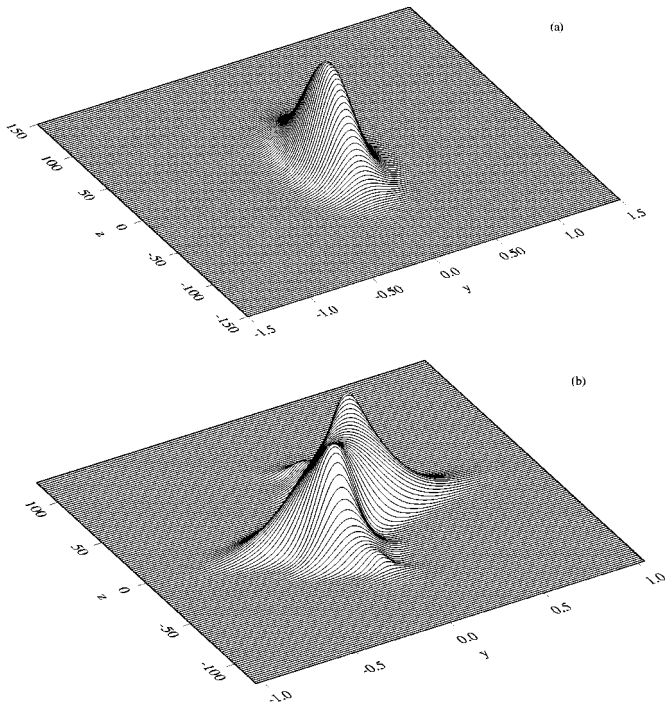

FIG. 5. Charged-particle case. The initial wave packet is shown in (a), and has spatial widths $\Delta y=0.13$ and $\Delta z\sim 20.7$. The parameter $\Omega^2=9.48$, and in (b) we show the spatial distribution of the wave packet at the end of the interaction time at $t=t_0=0.043$.

Equations (56) will now determine t_0 and Ω from Eqs. (76) and (87). For the values chosen above we will then obtain $t_0=0.043$ and $\Omega^2=9.48$.

D. Numerical results for a charged wave packet

With the chosen parameters, we integrate the Schrödinger equation (66) up to the time t_0. Figure 5(a) shows the initial spatial distribution, which is very narrow in the y direction [note the different scales on the axes of Fig. 5(a)]. The corresponding Gaussian momentum distribution is shown in Fig. 6(a). Figure 5(b) shows the spatial distribution of probability for the wave packet at the end of the interaction time t_0. The distribution is almost the same on both of the levels and so we only show one level here. The wave packet remains very narrow spatially, but its shape has been modified by a combination of spreading and twisting of different parts of the wave packet. Figure 6(b) shows the corresponding final momentum distribution. We can clearly see the splitting into two components which we expected from the arguments given in Sec. IV. The width in the z direction is small compared with the width in the y direction, but nevertheless the splitting can just be regarded as a spin separation.

Figure 7 shows the average forward velocity of the wave-packet components as a function of time. It is seen to decrease as time increases, because parts of the wave packet components are starting to turn back [especially if they are initially on the outside of the cyclotron line (79)]. However, at the time $t_0=0.043$ there still remains a net forward velocity of approximately unity in scaled units. In this example the transverse momenta at time t_0 are $p_y\sim 0.019$ and $p_z\sim\pm 0.043$, which means that the two spin components are emerging at an angle given approximately by (in scaled

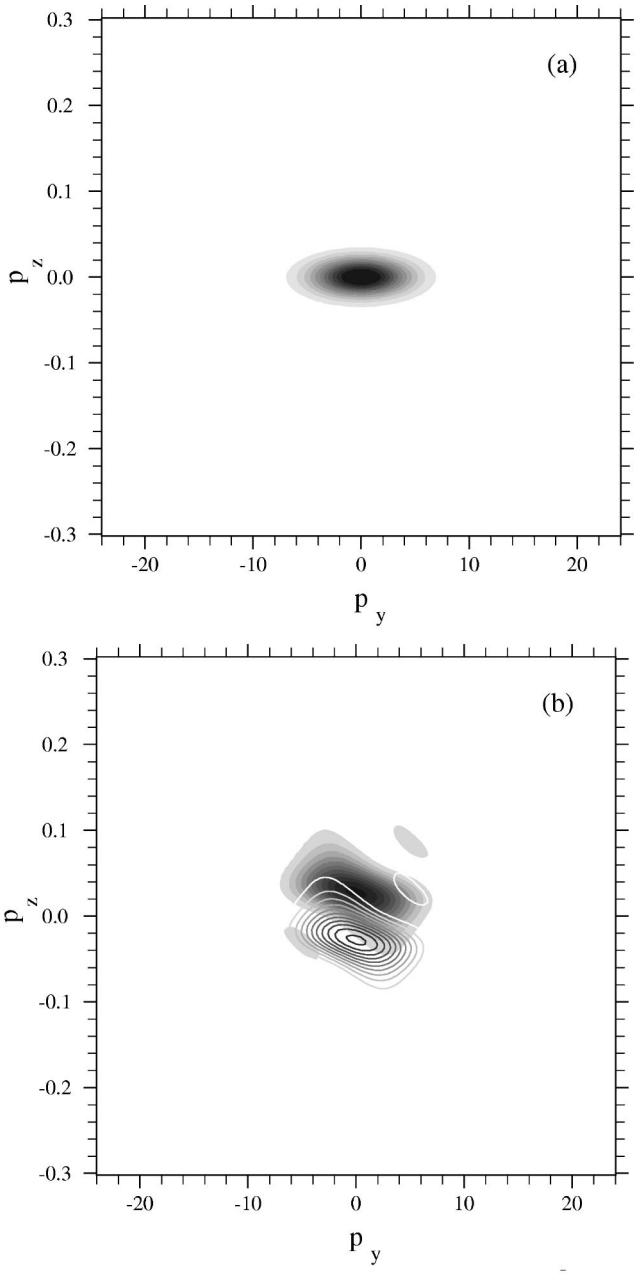

FIG. 6. Momentum distribution at (a) $t=0$, and (b) the scaled time $t=0.043$. Other parameters are as in Fig. 5. Both the wave-packet spin components are shown, the upper component being shaded to indicate height, while the lower component is marked with contour lines.

units) $\arctan(p_z/v_x)\sim 3°$ to the x axis. The crosses on the figure indicate the result from the scalar model of Sec. IV. For that model v_x is determined by Eq. (78) which leads to $v_x=\Omega^2/2-\xi^2+\eta^2$, where ξ and η are found from Eqs. (35). The scalar model provides a good fit for short times.

We may try to optimize the widths of the wave packet, the forward velocity, and the interaction time to improve the separation of the spin components. However, it appears to be difficult to achieve a substantially better splitting of the spins. The principal difficulty is that if the parameters reduce s, then there are many oscillations in the potential V_L [Eq. (69)], before a separation can take place; i.e., the angle $\Delta\phi$ [Eq. (77)], becomes large over the separation time scale.

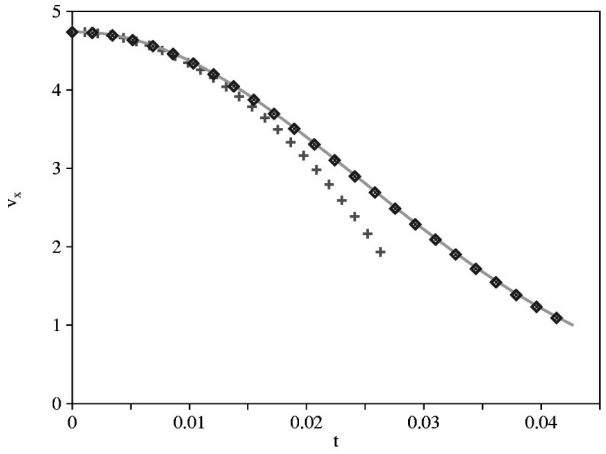

FIG. 7. The velocity v_x of one of the wave-packet spin components. The solid line shows the full quantum result computed from the wave functions and Eq. (78). The quantum result for the other spin component is extremely close to this curve. The crosses show the result from the scalar model of Sec. IV. The diamonds show the result from an ensemble of 50 000 classical simulations (Sec. VI). Parameters are as in Fig. 5.

Nevertheless, we will show in Sec. VII that we can utilize the potential V_L to obtain a substantial splitting, though it is in a different experimental configuration to that considered in this section.

VI. CLASSICAL SIMULATIONS

If we neglect the coupling between the two spin states, we can write down some simple classical equations based on the magnetic effects in the Newton equation:

$$m\frac{d}{dt}\mathbf{v}=-e\mathbf{v}\times\mathbf{B}\pm\mu\left(\frac{\partial B_z}{\partial z}\right)\hat{\mathbf{z}}. \tag{91}$$

If we use the quadrupole field (12) and the scaled units (32), we obtain

$$\frac{d}{dt}v_x=-2(zv_y+yv_z),$$

$$\frac{d}{dt}v_y=2zv_x, \tag{92}$$

$$\frac{d}{dt}v_z=2yv_x\pm 1.$$

As expected from Eq. (66), the quantity v_x+2yz is a constant of the motion. In order to match the quantum simulation, each member of the classical ensemble should have the same value of v_x+2yz. This means adjusting the initial velocity v_x according to the initial position.

We will now try to use a swarm of classical particles to replicate the dynamics of a wave packet. Figure 8(a) shows the initial momenta of such a swarm. The dispersion in position and momenta have been chosen so that the initial ensemble averages for the uncertainties match the quantum-mechanical ones. After the interaction time we obtain the swarms seen in Figs. 8(b) and 8(c), which can be directly

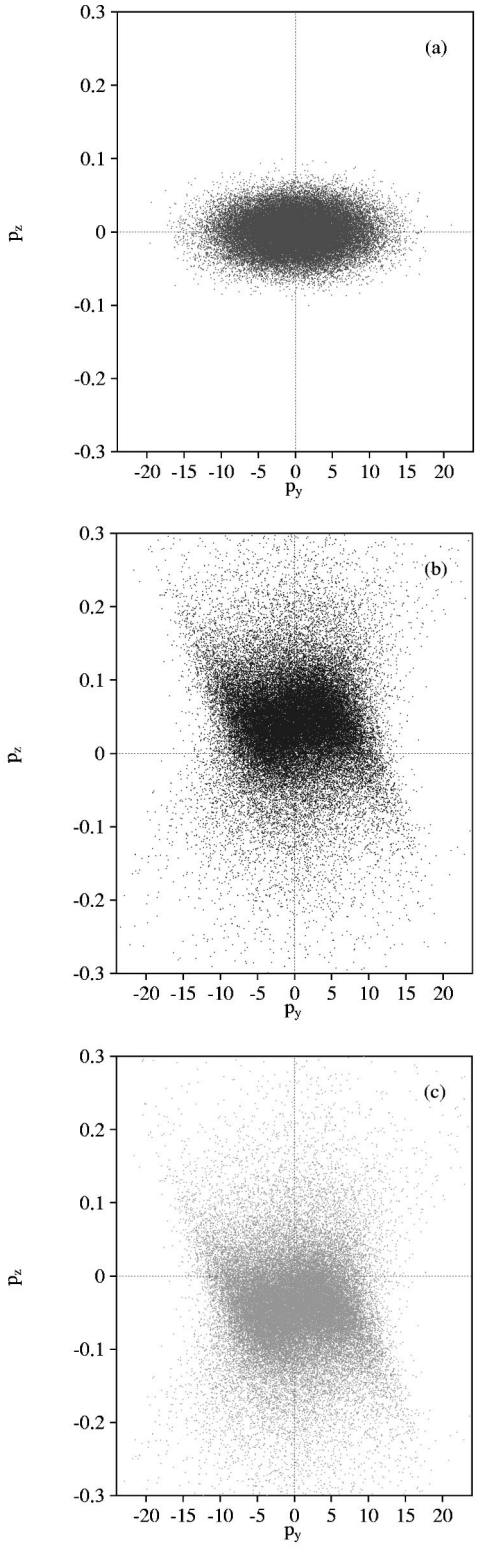

FIG. 8. Trajectories for classical particles in a simulation of the parameters in Fig. 5 with (a) the initial ensemble of momenta, and (b) and (c) the final ensemble of momenta at $t=0.043$. There are 50 000 trajectories in each figure.

compared with the quantum-mechanical result for the momentum distribution shown in Fig. 6. For the resulting splitting between the wave packets there is a good agreement between the classical and quantum results. In fact, the agreement between the ensemble averaged momentum and the

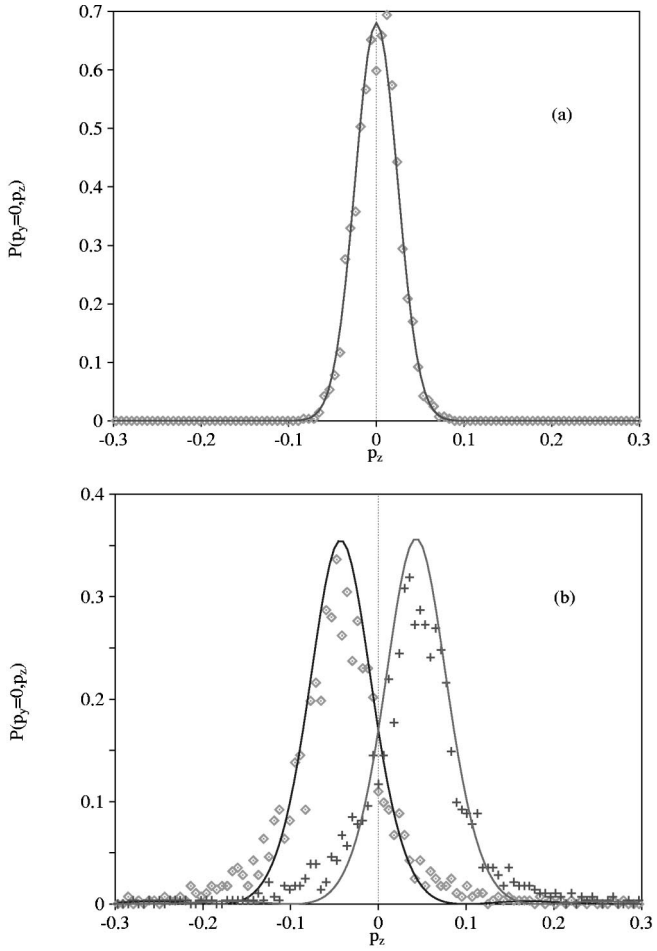

FIG. 9. Probability distributions of p_z for classical simulations (points) and quantum calculations (solid lines). We show a section through the 2D probability distribution at $p_y=0$. The classical ensemble result is taken from Fig. 8 by binning the data into 101 $\times$101 bins. Approximately 1900 trajectories are sampled from the ensemble for the figures. The quantum data are taken from sections of Fig. 6. In (a) we show the initial distributions of momenta, which are Gaussian, and in (b) we see the distributions at time $t=0.043$.

classical ensemble is excellent. For the classical simulations it is also straightforward to determine the forward velocity v_x. Figure 7 again shows excellent agreement between the classical and quantum results; this is despite the fact that the quantum result is indirectly calculated from the spatial wave functions by using Eq. (78). However, despite all this agreement, the shapes of the wave packets appear to be slightly different.

To examine the wave-packet shape more quantitatively we have taken a vertical section through the quantum data shown in Fig. 6 for $p_y=0$. The result is seen in Fig. 9, where the $t=0$ distribution is also shown. In the same figures, we show the corresponding distributions taken from the classical data. There is good agreement between the quantum and classical distributions at $t=0$ [Fig. 9(a)], as should be expected from the way the ensemble distribution has been generated. After the interaction time [Fig. 9(b)], we find good agreement between the quantum and classical values for the splitting of the wave packets, but it appears that the classical simulations lead to distributions of momenta that are too broad. Thus they do not display a resolution of the spin split-

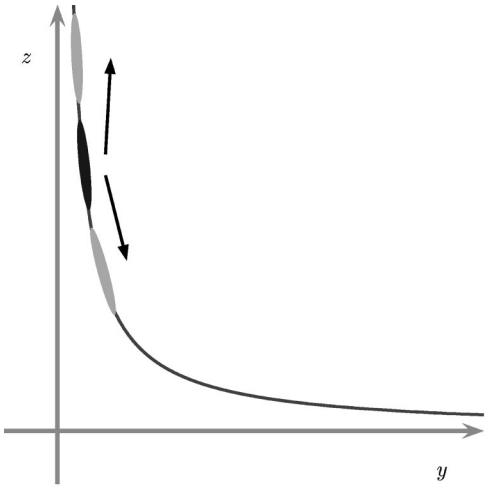

FIG. 10. This figure shows how we might expect a wave packet to split into spin components after being carefully placed on the cyclotron line. Here we show the expectation that the two components separate by moving along the cyclotron line in opposite directions.

ting which is as good as that of the quantum calculations. This suggests that it is important to perform the quantum calculations in this problem; classical simulations could be misleading concerning the issue of whether or not there is a significant amount of splitting of the spin components.

VII. BEATING THE LIMITS

A. Abiding with the Lorentz force

We can take a completely different approach to the spin-splitting experiment by abandoning the idea of the spin measurement as a beam experiment, and by trying to use the Lorentz force, which makes the spin splitting so difficult, to help the measurement. Our approach is to launch the wave packet along the cyclotron line so that it is in part constrained by the Lorentz potential (69). The general idea is illustrated in Fig. 10. The forces acting on the two spin components are still different, because the Stern-Gerlach force is not included in the Lorentz potential and thus the spin components can be separated. However, there are several disadvantages to this approach. First, the cyclotron line is curved, and does not lie on the z axis; consequently the measurement can only approximate a measurement of the z component of the spin. Second, if the wave packet is in the region of the cyclotron line, the velocity in the x direction, which was the beam direction, is close to zero. This derives from the fact that the cyclotron line is defined by the vanishing of the x velocity [Eq. (78)]. The electronic wave packet first has to be prepared appropriately and then subjected to the magnetic interaction in a controlled fashion.

B. Specification of the electron state

Our initial wave packet will again be taken to be a Gaussian of the form of Eq. (67), but with an initial velocity with components v_y and v_z. In order that the wave packet should travel smoothly in the Lorentz potential (69), any initial velocity has to be aligned along the tangent to the cyclotron line. From the fact that yz is constant on the cyclotron line, we will have $y\,\delta z + z\,\delta y = 0$, and hence the unit vector

$$\hat{e}_\perp = \frac{z}{r}\hat{e}_y + \frac{y}{r}\hat{e}_z \equiv \cos\theta \hat{e}_y + \sin\theta \hat{e}_z, \tag{93}$$

where $r^2 = y^2 + z^2$ is orthogonal to the tangent of the curve.

We now expand around a point $\{y_0, z_0\}$ on the cyclotron line, so that

$$y = y_0 + \delta y, \quad z = z_0 + \delta z, \tag{94}$$

and by writing

$$\delta y = \rho\cos\theta, \quad \delta z = \rho\sin\theta \tag{95}$$

we obtain

$$V(y,z) = 2(y_0\sin\theta + z_0\cos\theta)^2\rho^2 + O(\rho^3) \to 2(y_0^2 + z_0^2)^2\frac{\rho^2}{r^2}$$

$$= \frac{1}{2}k\rho^2, \tag{96}$$

with the spring constant

$$k = 4r^2. \tag{97}$$

This results in an angular frequency (in scaled units) of $\omega = \sqrt{k} = 2r$, in agreement with the cyclotron frequency (74) in scaled units. We will place our initial wave packet in the channel defined by Eq. (69), with a transverse width σ_T and an orientation which matches the width of the channel at the center of the packet. The choice of transverse width has to ensure that the center part of the wave packet does not breathe, i.e., periodically expand and contract, in the channel. If this were to happen, it would explore larger regions of the potential surface and the motion of the wave packet would cease to be usable for spin splitting. From the spring constant [Eq. (97)], we can determine the necessary transverse width to avoid breathing:

$$\sigma_T = \frac{1}{2\sqrt{r}}. \tag{98}$$

As mentioned earlier, the wave packet will be aligned along the local direction of the channel, which, however, will not remove the entire breathing effect because the channel spring constant varies slightly along the length of the wave packet.

For this configuration, we try to split the spin components spatially, which contrasts with the momentum splitting of the previous sections. In order to achieve this, it is essential that the wave packets separate by a distance greater than their width. The Lorentz forces tend to bend the wave packets along the cyclotron line, but if we assume that we are close enough to the z axis of the quadrupole field, we may assume that there is an approximate force of ± 1 (in scaled units) which splits the wave-packet components. This means that in a time t each wave packet will travel a distance $t^2/2$, and their separation will be approximately t^2.

During the time the wave packets are separating, they are also spreading in a longitudinal direction, that is, in a direc-

tion along the cyclotron line. This is because the wave packets are not constrained in this direction, though they are constrained in the transverse direction by the ''walls'' of the Lorentz potential. In free space the width of a wave packet increases as (see, e.g., Ref. [20])

$$\sigma(t)=\sigma_L\sqrt{1+\left(\frac{\hbar t}{2m\sigma_L^2}\right)^2}, \tag{99}$$

where σ_L is the initial longitudinal width of the wave packet. In terms of the scaled units (32), this is simply

$$\sigma(t)=\sigma_L\sqrt{1+\left(\frac{t}{2\sigma_L^2}\right)^2}. \tag{100}$$

If the wave-packet spin components are to separate by more than their width, we have to impose the condition

$$t^2>\sigma_L\sqrt{1+(t/2\sigma_L^2)^2}. \tag{101}$$

Then if we solve for t we find

$$t^2>\frac{1+\sqrt{1+(2\sigma_L)^6}}{2(2\sigma_L)^2}. \tag{102}$$

Minimizing this estimate gives the *shortest* possible time for spin separation as unity when $\sigma_L=1/\sqrt{2}$.

C. Numerical result

Figure 11(a) shows the initial wave-packet, which is highly unisotropic and orientated along the cyclotron line at some distance from the origin. As the wave packet dynamics proceed, the packet splits into two components which are well separated as can be seen at $t=3.4$ in Figs. 11(b) and 11(c). Equation (102) predicts that $t=2.08$ would be sufficient, given the parameters of Fig. 11, to separate the spin components, but by examining the system at a later time we can find improved spin separation.

Figure 11(b) shows some traces of the wave packet seen in Fig. 11(c). Presumably this is because the wave packet is, initially, tilted away from the z axis, by an angle of 0.117°, which results in a slight admixture of the two spin components. If the distance from the origin were to be increased, this type of effect would decrease. However, it would happen at the expense of the required transverse width becoming increasingly small.

A careful inspection of Figs. 11(a) and 11(b) also shows that one final wave-packet component has hardly changed its position as compared to where it was initially. This seems strange when there was supposed to be a force on this component of $+1$ (in scaled units) from the interaction with the gradient of the magnetic field. This force is almost directly along the channel of V_L. Likewise, the wave packet in Fig. 11(c) has been subjected to double the expected acceleration.

It is possible to give a straightforward explanation of this effect by using some simple quantum mechanics. Along the cyclotron line, at a distance r from the origin, the potential V_L appears in the transverse direction of the cyclotron line as an approximate harmonic potential with (as described above) Eq. (97), a spring constant of $k=4r^2$. The wave packet is

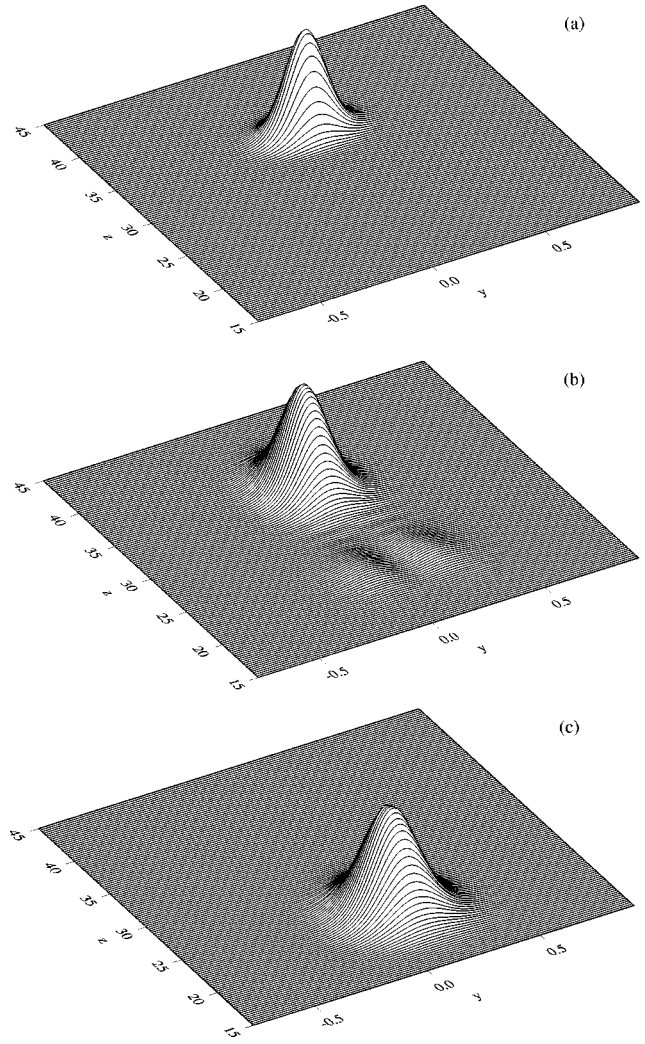

FIG. 11. A wave packet is initially placed on the cyclotron line (a), and it subsequently splits into the two very separate components seen in (b) and (c). The wave packet is initially placed at $y_0=0.0714$ and $z_0=35.0$ with a transverse width $\sigma_T=0.0845$ as determined by Eq. (98). We have chosen a longitudinal width of $\sigma_L=1$ and $\Omega^2=9.48$. The wave packets in (b) and (c) are shown at the scaled time $t=3.40$ (chosen to be sufficient to achieve good separation).

designed to have its transverse width such that it is in the ground state of the transverse potential, and thus it can avoid any breathing motion. However, if we move the wave packet along the cyclotron line, toward or away from the origin, the spring constant will change, which will change the ground-state energy. The ground-state energy is, from Eq. (97), $E_0=\omega/2=\sqrt{k}/2=r$, and if the splitting force tries to move the wave packet away from the origin, the increase in transverse ground-state energy will act to oppose the motion. Conversely, if we move in a direction approaching the origin, the decrease in transverse ground-state energy will assist the wave-packet motion, provided that the wave packet remains in the transverse ground state, i.e., remains adiabatic. We can estimate the size of the effect from the transverse ground-state energy. For small y, $r\sim z$ and the ''force'' from the change in ground-state energy will almost exactly cancel the splitting force in the $+z$ direction. The result is that a change

in the transverse ground-state energy strongly affects the longitudinal motion along the cyclotron line.

VIII. SUMMARY

In this paper we have tried to address the problem of the separation of the spin components of an electron, from a point of view that is in the spirit of the original argument attributed to Bohr and Pauli; only Sec. VII of the paper presents an alternative approach. This primarily assumes some kind of electron-beam device, utilized to separate the spin components. We have shown that there may be some flexibility in the physics of the scheme, which is introduced by allowing the electron beam to be unisotropic in cross section. This approach appears not to have been considered in the original Bohr argument as reported by others [5,6]. We have developed the idea by introducing a scalar model of the system [Eq. (30)], which neglects two dynamical effects which are also not included in the original argument. The two neglected effects are the off-diagonal effects from the spin coupling, i.e., the y term in Eq. (13), and the anharmonic effects from the $\mathbf{A}^2$ term in the Hamiltonian. The scalar model is solvable exactly, and we use the properties of the solutions to suggest parameters where a spin splitting might be possible and to guide us in more precise and fully quantum-mechanical, though numerical, calculations.

The numerical approach allows the inclusion of the two effects omitted in the scalar model, and a testing of the effectiveness of the spin splitting in the system. We find, as may be seen in Fig. 6, that a spin splitting can exist, although there is still some overlap of the spin components. Nevertheless, a statistical analysis of the resulting spin distribution, such as seen in Fig. 9, could be used as the basis of a spin measurement. A complication is the bending of the electron beam in the magnetic field, because of the Lorentz force (i.e., the cyclotron motion). However, we are able to determine the beam velocity of the wave packet, and this confirms that there is forward motion even when the spin components are split. Forward motion implies a beamlike behavior, and in this sense we feel that we have illustrated what can be done, to stretch the situation of the original argument beyond its limit by modifying the experimental configuration. We are at its limits of performance because, even in our setup, where the spins are split in momentum space in the z direction, the electron beam becomes very spread out in the y direction, and the velocity of the electrons in the x direction is rather reduced. The fact that it is not easy to separate the spins, and that there is still an overlap of the spin components, is one testament to the insight of Bohr. We, however, conclude that the spin measurement is more a matter of experimental skill than one of fundamental issues. In principle, our argument applies to all charged elementary particles having a magnetic moment of the type (5). However, with the proton, for example, the g factor is larger than that of the electron; this acts to improve the separability of the spin states.

So where does the approach of Bohr and Pauli differ from our treatment? One way is their promotion of a practical complication to a fundamental principle. Furthermore their argument is based on using p_x/m as the velocity in the Lorentz force expression. The actual physics demands the use of the velocity v_x itself, and, situating the initial wave packet along the z axis, we make large parts of it lie very close to the potential minima we have called the cyclotron lines. Here the actual velocity is very small, which we surmise may be the origin of the improvement we have found in the spin resolution; the average Lorentz force is not as efficient in degrading the result, as suggested by the original argument. One may also say that Bohr and Pauli did not treat the full quantum problem, even though quantum mechanics enters the problem as an uncertainty.

Our approach can be criticized on many fronts, and these issues form the basis of sufficient discussion to warrant further investigations of the details. It can be argued that we have not properly calculated the effect on the wave packet of entering the magnetic field. This is because we have made a model 2D calculation which does not include variations, e.g., in the magnetic field, along the beam direction. There are two physical approaches to this: a sudden application of the magnetic field, or a smooth, adiabatic turn-on. We intend to make further studies of the effects of the slow turn-on of the field in more realistic calculations. Likewise, we have only considered a model quadrupole field. Real magnetic fields are different, and it is not impossible that some other field arrangement might yield improved splitting effects. Furthermore, by shaping the electron beam into a non-Gaussian profile, it may be possible to optimize the splitting further.

We can also be criticized for not presenting a detailed model of the full detection process. The numerical calculations provide only a splitting in momentum space. Of course, this is enough, in principle, to split the spin states. Also, we have shown that in subsequent evolution the two spin components will separate in real space. An alternative approach would be to use an electron lens to transform the momentum-space splitting to a real-space splitting. After the interaction with the field, part of the electron wave packet moves in a backward direction. Because the spin separation is achieved in momentum space, it is not clear at this time how an observed splitting might be affected by any correlation between the momentum components and spatial location in the beam direction. However, the cuts at $p_y=0$ in Fig. 9, essentially show the measurement result from an ideal flat detector in the x-z plane, where the beam is expected to display the spin splitting. All parts of the wave packet having a positive v_x component will reach such a detector. There is no need to assume that the detector will cover all values of p_y, even if a real one, of course, integrates over a finite slice in space.

It was not our intention in this paper to discuss details for a potential experiment, as our main emphasis has been on discussing the validity of a point of principle. However, it may well be of interest to examine the values of real parameters compatible with the scaled values chosen for Fig. 6. We take an electron with velocity v_x, and make an estimate of l by using v_x to determine λ_x [Eq. (31)] and $\Omega^2=9.48$ (from Figs. 5 and 6). We then can determine the real wave-packet sizes from the scaled ones, and can also calculate the field gradient required from Eqs. (32). Then, for example, with $v_x\sim 100$ m/s we find that, for the parameters in Fig. 6, the wave-packet widths are 0.7 μm and approximately 110 μm. The required field gradient is then 8 T/m over millimeter distances. This is a rather low-energy electron, only 0.03 μeV, but the energy could be increased if smaller electron

wave packets can be prepared. For example, a reduction in the linear dimensions of the wave packet by a factor of 10 results in a hundredfold increase in the kinetic energy of the electron. Despite these results for $\Omega^2 = 9.48$, we add that we have not tried to optimize the parameters from an experimental point of view; i.e., we have been partly guided by computational constraints, and the possibility of more optimal parameters for an actual experiment is not necessarily ruled out.

Clearly there remain many issues which it would be of interest to resolve. This is especially true when one considers experimental consequences of our theoretical considerations. We hope that in this paper we have shown that there is some merit in considering these problems more deeply.

Finally, we turn to the argument presented in Sec. VII. This is not a beam arrangement, because there is no net forward motion of the electron, and in that sense it is not within the realm of what we know of the Bohr discussion. However, it is interesting as a method for providing a clear spatial separation of electron-spin components by arranging for the wave-packet components to reside in a quantum-mechanical ground state of the parts of the Lorentz force field that do the worst damage in the effort to split the spin components. The direction of motion of the wave packets is close to the direction of the gradient of the field, and in that sense there is a resemblance to the arrangement considered by Brillouin [21] (where the classical electrons approach a magnetic pole) and the classical calculations of Ref. [12].

In summary, Bohr's argument, which appears to have been that it is not possible to separate the spin components of the electron *in principle*, appears to be refuted, even within the spirit of the original argument. We have made quantum-mechanical calculations in a simplified model which are indicative of such a result, and numerical calculations of the quantum-mechanical behavior which support the conclusions of our analytic model.

ACKNOWLEDGMENTS

S.S. would like to acknowledge discussions with Sir Rudolph Peierls. B.M.G. would like to thank the Wenner-Gren Foundation for financial support, and KTH for hospitality in Stockholm. We would also like to thank the Niels Bohr Archive for Ref. [13].

[1] M. Nauenberg, C. Stroud, and J. Yeazell, Sci. Am. **270** (6), 44 (1994); M. Nauenberg, Phys. Rev. A **40**, 1133 (1989).
[2] N. Bohr, J. Chem. Soc. **134**, 349 (1932); **134**, 368 (1932).
[3] N. Bohr, Nature (London) **121**, 580 (1928).
[4] *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, 1983), p. 699.
[5] W. Pauli, in *Le Magnetisme*, Proceedings of the VI Solvay Conference (Gauthier-Villars, Paris, 1932), Sec. II.1.b, p. 175.
[6] N. F. Mott, Proc. R. Soc. London, Ser. A **124**, 440 (1929).
[7] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions*, 3rd ed. (Clarendon Press, Oxford, 1965), pp. 214–219.
[8] W. Pauli, in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1958), Vol. 5, Pt. 1, p. 165.
[9] See, for example, J. Dupont-Roc, C. Fabre, and C. Cohen-Tanoudji, J. Phys. B **11**, 562 (1978).
[10] R. S. van Dyck, P. B. Schwinberg, and H. G. Dehmelt, Phys. Rev. D **34**, 722 (1986).
[11] Sir Rudolf Peierls (private communication).
[12] H. Batelaan, T. J. Gay, and J. J. Schwendiman, Phys. Rev. Lett. **79**, 4517 (1997).
[13] Letter from N. Bohr to N. F. Mott, dated March 1929 (see Bohr Scientific Correspondence Microfiche No. 14). The critical passage reads ''Since you left, I have sometimes thought of the problem of the realisation of the electron polarisation, and after all I am quite prepared that such a polarisation might be observable. In fact the argument I used in the discussion of the Stern Gerlach effect was not strict due to the mixture of classical mechanics and wave theory in the very region where no sharp distinction is possible. Thus the argument tells only that as regards the electron magnetization we cannot as in the ordinary discussion of the Stern-Gerlach effect base our considerations on classical pictures of a moving magnet, but not that the closer quantum-theoretical treatment will never give a positive effect.''
[14] G. H. Rutherford and R. Grobe, Phys. Rev. Lett. **81**, 4772 (1998); H. Batelaan and T. J. Gay, *ibid.* **81**, 4773 (1998).
[15] G. H. Rutherford and R. Grobe, J. Phys. A **31**, 9331 (1998).
[16] F. Adler, Helv. Phys. Acta **10**, 455 (1937).
[17] S. Stenholm, in *Quantum Measurements in Optics*, edited by P. Tombesi and D. F. Walls (Plenum, New York, 1992), p. 15.
[18] See, for example, A. Carnegie and I. C. Percival, J. Phys. A **17**, 801 (1984).
[19] E. Enga and M. Bloom, Can. J. Phys. **48**, 2466 (1970).
[20] B. M. Garraway and K.-A. Suominen, Rep. Prog. Phys. **58**, 365 (1995).
[21] L. Brillouin, C. R. Hebd. Seances Acad. Sci. **184**, 82 (1927); Proc. Natl. Acad. Sci. USA **14**, 775 (1928).